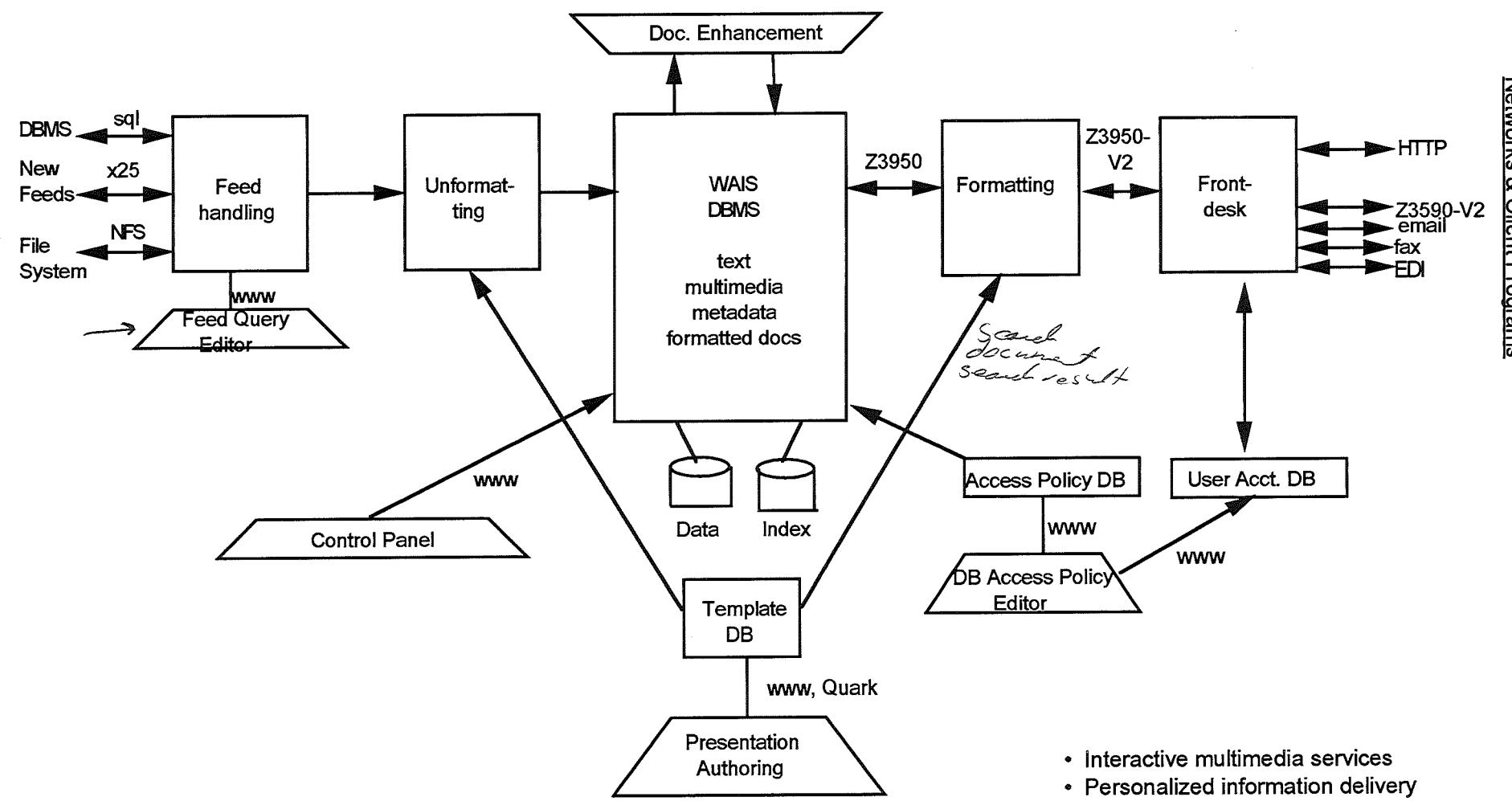


INTERNET PUBLISHING PLANT

7/18/94

Company Confidential



- Interactive multimedia services
- Personalized information delivery
- Turn-key operation for publishers

WAIS Technology Architecture

Brewster Kahle

7/12/94

DRAFT

Company Confidential

This document is the start of the next generation WAIS technology architecture that will be used for our in-house production services and our technology products. As a complete publishing system WAIS Inc will be able to deliver parts or all to customers. Parts of this will be licensed from other vendors.

Contents:

- WAIS Technology Architecture diagram
- Component Functional Overview
- Architectural Considerations
- In-house development plan <to come>
- Partnering needs <to come>

Component Functional Overview

Feed Handlers: Automated query generators
Ad hoc libraries
Control panel?
query editor?
x.25, SQL, IP email (Rainman), news feeds, NFS

Unformatting: (currently the custom parser? toolkit)
Finds fields, metadata, etc.
Formats this data for the DBMS
Can this be controlled easily? parse by example?

Doc Enhancement:
Triggered by DB additions, periodic queries, or human interaction
Adding metadata, related documents, hypertext links, keywords

WAIS DBMS: Text, multimedia, metadata
Stores structured documents
Points to original formatted documents
Indexes documents
Add, delete, modify, (commit)(commit to staging area)?

Formatting: HTML creation for documents, search results, browse results
Uses templates to decide how to do it; which templates comes from DB & doc records?
Talks structured protocols (I think)

Front desk: User handling, welcome-screens, help-screens
Authentication of users & superusers
New user screens, user account screens
Profile storage, agent handling
HTTP, Z3950-U2, email, fax

User Acct. DB: Current status, logged changes in status
Profiles & profile execution logs
Preferences & personalization features
Group membership list (enough for access restrictions?)
Group preferences & personalization

DB access policy editor: How do we express db policies?
What's a db?

Group access issues

Access Policy DB: Stores DB access policies

Presentation Authoring: Defining HTML templates
Can it be used for creating parsing templates?

Template DB: Storage

Control Panel: Monitors server use
Deletion, restoring, modify?
Shutdown

Metadata: Category codes
Links to related documents
Author
Date
source
Headlines
Formats, original URL's
Cost
Past usage statistics
Permission

Architectural Considerations

- Notes integration
- BBoard/conf. system handling
- Agent interaction
- Fax delivery
- CD ROM productions
- Video clips
- Help desk
- Training
- Catalog application
- Annotations
- Personalization of screen
- External db integration
- Can all actions be taken remotely? all through WWW? remote edition of a doc or script?

 Adding docs

 Adding templates

 Managing collections

- Pre-press integration
- Per article tweaking of presentation
- Arbor Software
- EDI
- Compare & contrast

 Sybase, Oracle, Mosaic Inc., library system

In-house Development Plan

Partnering needs